

**Armin Baur**

*Pädagogische Hochschule Schwäbisch Gmünd*

## **Inwieweit eignen sich bisherige Diagnoseverfahren des Bereichs Experimentieren für die Schulpraxis?**

**To what extent are present diagnosis procedures of the area of experimentation suitable for use in school?**

Eine Individualisierung des Unterrichts setzt eine gute Diagnostik voraus. Die Entwicklung und Bereitstellung von schulpraxistauglichen Diagnoseinstrumenten (Pädagogische Diagnostik) ist daher eine wichtige Aufgabe. Im Bereich Experimentieren gibt es viele Arbeiten, die sich mit wissenschaftlicher Kompetenzdiagnostik auseinandersetzen. In diesem Artikel werden vorhandene wissenschaftliche Diagnoseverfahren für den Bereich Experimentieren (Realtests, Paper-pencil-Tests, computerbasierte Testumgebungen und Interviews) dargestellt und auf ihre Einsetzbarkeit für die Pädagogische Diagnostik diskutiert. Als geeignet erscheinen vom Lernenden angefertigte Protokolle und partielle Beobachtungen von Schüler/innen. Weiterhin wird die Notwendigkeit eines Theorierahmens angeführt, der Experimentierkompetenz für die Pädagogische Diagnostik in diagnostischen Analyseeinheiten (Teilaspekten) beschreibt und Vergleichswerte bietet.

**Schlüsselwörter:** Diagnose, Diagnoseverfahren, Individualisierung, Differenzierung, Experimentierkompetenz

Individualisation in teaching imperatively necessitates good pedagogical diagnostics. It is an important task for educational science as well as for science education to create instruments for diagnosis which can be used in school. Concerning aspects of experimentation, there has already been done much research on competence diagnostics. In this article, existing diagnosis methods regarding experimentation (performance test, paper-pencil-tests, computer simulations, and interviews) are juxtaposed and analysed for applicability in school (pedagogical diagnostics). Protocols drawn up by pupils and cursory/exemplary observations of pupils appear to be applicable. Furthermore, the necessity of a theoretical framework is stressed, which is to describe experimentation competence in analytical chunks (partial aspects) applicable in pedagogical diagnostics and which can arrange empirical data in a frame of reference.

**Keywords:** diagnosis, diagnosis methods, individualisation, differentiation, experimentation competence

## 1 Hintergrund

Die Notwendigkeit der Beachtung der Heterogenität einer Lerngruppe ist in der aktuellen Pädagogik unbestritten (Buholzer & Kummer Wyss, 2010, S. 8). Eine Berücksichtigung führt zu einer Individualisierung des Unterrichts (durch Eingehen auf den individuellen Schüler/die individuelle Schülerin) und dies wiederum zur Differenzierung. Diversität und Binnendifferenzierung bilden nicht nur wichtige Aspekte innerhalb der Schulpädagogik, sondern auch innerhalb der Fachdidaktiken der Naturwissenschaften (siehe hierzu auch Abels & Markic, 2013; Krüger & Meyfarth, 2009; Pysik & Bauer, 2013). Neigungs- und Leistungsdifferenzierte Bildungsangebote werden explizit auch von der Kultusministerkonferenz für den MINT-Bereich empfohlen (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland [KMK], 2009). Eine individuelle Förderung von Schüler/innen setzt wiederum eine professionelle und sachgemäße Pädagogische Diagnostik voraus (Hesse, 2014).

*Pädagogische Diagnostik* beinhaltet sämtliche diagnostischen Tätigkeiten, über die bei Lernenden die Voraussetzungen und Bedingungen planmäßiger Lehr- und Lernprozesse ermittelt, Lernprozesse analysiert und Lernergebnisse festgestellt werden, damit individuelles Lernen optimiert werden kann. Zur Pädagogischen Diagnostik gehören zudem auch Tätigkeiten, die die Zuweisung zu Lerngruppen oder zu individuellen Förderprogrammen ermöglichen. Gleichmaßen gehören Tätigkeiten dazu, die auf das Erteilen von Qualifikationen zielen. (Ingenkamp & Lissmann, 2008, S. 13)

Pädagogische Diagnostik grenzt sich von der wissenschaftlichen Diagnostik ab. Die Pädagogische Diagnostik ist nicht auf allgemeine Zusammenhänge ausgerichtet, sondern auf den Einzelfall (Klauer, 1982). Im formalen Bereich kann eine Pädagogische Diagnose nur unter Einbeziehung geeigneter Methoden (Diagnoseinstrumente und -verfahren) erfolgen. Eine methodische Fundierung ist mit einer theoretischen Fundierung verbunden (Ingenkamp & Lissmann, 2008; Tent & Stelzl, 1993). Tent und Stelzl (1993) führen an, dass eine Gegenstandsbestimmung und eine theoretische Einordnung dessen, was diagnostiziert werden soll, notwendig sind. Hierzu bietet sich nach Erachten des Autors ein Theorierahmen an, der:

- (1) einen Überblick von bei Schüler/innen vorkommenden Lernschwierigkeiten (Barrieren, die eine Ausbildung bzw. Weiterentwicklung von Teilkompetenzen verhindern) bietet, mit dem Lehrkräfte Lernende diagnostizieren können (siehe hierzu Aldersen, 2005);
- (2) die Möglichkeit eröffnet, Lernende bezüglich einer altersgleichen Vergleichsgruppe in ihrem Leistungsstand anhand von Vergleichswerten einzuschätzen (Hesse & Latzko, 2011), um individuelle Lernvoraussetzungen zu erkennen und dadurch sinnvoll Fördermaßnahmen (Übungen etc.) abzuleiten. Vergleichswerte können hier Lernergebnisse oder auftretende Fehlvorstellungen und Fehler sein.

Erst mit solch einem Theorierahmen und darauf abgestimmten Diagnoseinstrumenten und -verfahren können aussagekräftige Diagnosebilder erstellt werden. Innerhalb der Naturwissenschaftsdidaktiken setzt man sich bereits intensiv mit der Kompetenzdiagnostik auseinander. Aber wie geeignet sind unter dem Blickwinkel der Schulpraxis die entwickelten Instrumente und Verfahren zur wissenschaftlichen Kompetenzdiagnostik für eine

Pädagogische Diagnostik? Reicht die Entwicklung von Diagnoseinstrumenten und -verfahren für die Schulpraxis ohne eine wissenschaftliche Weiterentwicklung, Präzisierung und Prüfung von diagnostischen Analyseeinheiten (Teilkompetenzen, Fehler von Schüler/innen etc.) zur Erfassung von Lernschwierigkeiten bzw. des Lernstandes aus?

Unter den genannten diagnostischen Analyseeinheiten sind definierte Beurteilungskategorien zu verstehen, die erfasst werden und ein Diagnosebild erzeugen.

In dieser Arbeit werden beide Fragestellungen für den Kompetenzbereich Erkenntnisgewinnung und hier speziell für das Experimentieren aufgegriffen und beantwortet. Experimentieren wird als Problemlöseprozess zur Erkenntnisgewinnung verstanden (vgl. hierzu Klahr, 2000; Mayer, 2007; Wellnitz et al., 2012), der von den Schüler/innen selbstständig durchgeführt wird - im Sinne des offenen Experimentierens. Bei der Verwendung der Begrifflichkeit „offenes Experimentieren“ sind in diesem Artikel alle von Priemer (2011) beschriebenen Offenheitsdimensionen einbezogen.

## **2 Anforderungen an Diagnoseinstrumente und -verfahren aus schulpraktischer Sicht**

Lernverlaufs- und Lernstandsdiagnostik sind Grundlage für eine richtige Beratung, Bewertung und Förderung eines Lernenden sowie zur Erteilung von Qualifikationen (Hesse, 2014). Ein Fehlurteil kann gravierende Konsequenzen nach sich ziehen. Aus diesem Grund muss „[...] die Tätigkeit des Unterscheidens und Urteilens in einem professionellen Kontext nach einem theoretisch begründbarem System von Regeln und Methoden zur Gewinnung und Analyse von Daten ablaufen.“ (Hesse & Latzko, 2011, S. 59). An dieser Stelle ist wichtig anzuführen, dass einige Autoren bewusst zwischen Diagnose und Leistungsmessung im Sinne der Erstellung von Qualifikationen unterscheiden (z.B. Huhta, 2008; Langlet, 2009). Die Position von Langfeldt (2006) vermittelt zwischen den unterschiedlichen Ansichten, sodass Pädagogische Diagnose gleichzeitig drei Ebenen anspricht:

- *Individuelle Ebene:* Auf dieser Ebene müssen von der Lehrkraft Lernvoraussetzungen und Leistungsmöglichkeiten der einzelnen Schüler/innen erkannt werden.
- *Klassenebene:* Auf der Klassenebene muss die Lehrkraft z. B. die interindividuellen Leistungsunterschiede in einer Klasse erkennen können, um bei kooperativen Lernformen geeignete Gruppen zusammenzustellen, oder sie diagnostiziert die Interessenslage und Präkonzepte der Klasse, um geeignete Themenbereiche für den Unterricht auszuwählen (Langlet, 2009).
- *Institutionelle Ebene:* Auf der institutionellen Ebene muss die Lehrkraft gerechte und objektive Leistungsbewertungen (z. B. Zeugnisse) vollbringen.

In dieser Arbeit wird der Blick auf die individuelle Ebene gerichtet. Der formellen Pädagogischen Diagnostik, die zielgerichtet, theoriegeleitet mit wissenschaftlich geprüften Methoden arbeitet, stehen informelle subjektive Urteile gegenüber (Schrader, 2010): „Während formelle (d. h. auf erprobte diagnostische Methoden gestützte) Diagnosen immer dann wichtig sind, wenn es um langfristige und für den Betroffenen bedeutsame erzieherische

Entscheidungen und Maßnahmen geht, sind informelle Urteile für die kurzfristige Steuerung des erzieherischen Handelns unverzichtbar“ (ebd., S. 107).

Damit formelle Diagnostik sich vom subjektiven Urteil unterscheidet, müssen Diagnoseinstrumente und -verfahren (Diagnosetests, Diagnosegespräche und diagnostische Beobachtung) für die Arbeit auf der individuellen Ebene folgenden schulpraktischen Anforderungen (SAs) standhalten:

(SA1) „Erfüllung der Testgütekriterien“:

Diagnoseinstrumente und -verfahren müssen die Testgütekriterien (Objektivität, Reliabilität, Validität) erfüllen (Schrader, 2010; Walter, 2008).

(SA2) „Theoretische Präzisierung“:

Zu messende Merkmale müssen theoretisch präzisiert sein (Hesse & Latzko, 2011). Da die Ergebnisse (Aussagen) der Diagnose für die Beratung und Förderung eingesetzt werden, muss unmissverständlich (ohne Interpretationsspielraum) sein, was gemessen wurde.

(SA3) „Lernstand sichtbar machen“:

Formelle Diagnoseverfahren müssen Entwicklungsstände, Lernpotentiale bzw. Lernhindernisse sichtbar machen (Alderson, 2005). Diese Aspekte der Diagnose sind für die individuelle Ebene unabdingbar und von Lehrkräften wird erwartet, dass sie darum wissen (Standards für die Lehrerbildung, KMK, 2004). Das sachliche und objektive Erkennen dieser Aspekte kann aber sicherlich nicht umfassend durch alleiniges alltägliches Beobachten und Urteilen (informell) erfolgen. Geeignet erscheint eine Kombination von alltäglichem Beobachten und einer Erfassung mittels geeigneter Diagnoseverfahren (formell).

(SA4) „Veränderungen sensibel erfassen“:

Formelle Diagnoseverfahren müssen auch schon kleine Veränderungen des Lernstandes erfassen, damit eine individuelle Förderung des Lernenden stattfinden kann (Walter, 2008).

(SA5) „Mehrfach einsetzbar“:

Formelle Diagnoseverfahren müssen mehrfach einsetzbar sein, um eine Veränderung sichtbar zu machen (ebd.).

(SA6) „Vorhandene Beurteilungsmaßstäbe“:

Es müssen Beurteilungsmaßstäbe zur Interpretation der Messergebnisse vorhanden sein (Hesse & Latzko, 2011). Für Beratungsgespräche und Fördermaßnahmen ist die Einordnung eines Schülers/einer Schülerin gegenüber einer Vergleichsgruppe und einer kriterialen Bezugsnorm (Orientierung an den Lernzielen) wichtig, um zu entscheiden, in welchem Ausmaß gefördert werden muss.

(SA7) „Zeitökonomie“:

Der Einsatz und die Auswertung von formellen Diagnoseverfahren müssen unkompliziert und zeitökonomisch sein (Walter, 2008).

(SA8) „Für Rückmeldung geeignet“:

Die erfassten Ergebnisse müssen für die Mitteilung/Information an Schüler/innen und Eltern geeignet sein, da erst durch eine Rückmeldung der formellen Diagnose pädagogische Wirkung erzielt wird (Ingenkamp & Lissmann, 2008).

### **3 Bisher entwickelte/verwendete Diagnoseinstrumente und -verfahren zur wissenschaftlichen Erhebung von Experimentierkompetenz**

Nach Schreiber, Theyßen und Schecker (2009) können drei Verfahren zur Erhebung experimenteller Kompetenz unterschieden werden: schriftliche Wissenstests, Experimentaltests und computerbasierte Testumgebungen. In der Literatur findet man zusätzlich noch die Verwendung von Interviews, was zu vier möglichen Verfahrenstypen führt, die in manchen Untersuchungen auch trianguliert verwendet werden. Die vier Verfahrenstypen können nach dem Grad der benötigten kognitiven sowie manuellen Fähigkeiten und Fertigkeiten drei unterschiedlichen Realitätsdimensionen zugeordnet werden (reales Experimentieren, virtuelles/gedankliches Experimentieren, Nachdenken über Experimentieren).

(I) „Reales Experimentieren“:

- *Experimentaltest mit Realexperiment (performance assessment, practical test, laboratory performance test)*:

Die Verwendung von Experimentaltests ist die Methode, die in der Literatur häufig zur Diagnose experimenteller Kompetenz angesprochen wird. Bei Experimentaltests führen Schüler/innen ein reales Experiment durch. Die Daten werden je nach Forschergruppe entweder durch gezielte direkte Beobachtung (Kirchner & Priemer, 2007), Videographie und kriteriengeleitete Auswertung (Meier & Mayer, 2012; Wahser & Sumfleth, 2008), Auswertung von während des Experimentierens angefertigten Schüler/innen-Protokollen (Emden & Sumfleth, 2012; Kirchner & Priemer, 2007) oder anschließenden Interviews mit kriteriengeleiteter Auswertung (Schauble, Klopfer & Raghavan, 1991) ermittelt.

(II) „Virtuelles/gedankliches Experimentieren“:

- *Schriftlicher Test (Paper-pencil-Test)*:

Hier sind in der Literatur zwei verschiedene Methoden zu finden: Multiple-Choice-Aufgaben und Planen von gedanklichen Experimenten.

Multiple-Choice-Aufgaben: Hammann, Hoi und Bayrhuber (2007) konzipieren auf der Grundlage des SDDS-Modells (Scientific Discovery as Dual Search-Modells) Multiple-Choice-Aufgaben zum Experimentieren, mit denen die Experimentierkompetenz von Schüler/innen gemessen werden kann. Das Messwerkzeug ist reliabel und weist zwei Dimensionen („Testen von Hypothesen“ und „Suche im Hypothesenraum und Analyse von Evidenz“) auf. Eine Erweiterung und Weiterentwicklung des Erhebungsbogens erfolgte von Schulz, Prinz und Wirtz (2012). Der von ihnen entwickelte Bogen ermöglicht, dass Proband/innen Experimente durch Auswählen (Ankreuzen) von Variablen planen und nicht nur wie bei Hammann et al. (2007) die Antwort aus vier Alternativen wählen.

Planen eines Gedankenexperiments: Hammann, Hoi, Ehmer und Grimm (2008) verwenden bei einer Untersuchung zusätzlich zu ihren Multiple-Choice-Aufgaben eine Aufgabe, bei der

Schüler/innen ein Experiment gedanklich planen und dieses aufschreiben müssen. Die geplanten Experimente der Schüler/innen wurden später kriteriengeleitet ausgewertet.

*- Computerbasierte Testumgebung:*

Dieses Verfahren wurde u. a. von Schreiber, Theyßen und Schecker verwendet und in Bezug auf die Kompetenzdiagnose beim Experimentieren hin untersucht. Die Schüler/innen führen hierbei Experimente in einer Computersimulation durch und werten die resultierenden (computeranimierten) Ergebnisse aus (Schreiber, Theyßen & Schecker, 2009).

*- Interview:*

In der Literatur ist auch die Methode des Interviews, welches das Planen und Auswerten eines gedachten Experimentes dokumentiert, zu finden. Schrempp und Sodian (1999) verwenden eine Geschichte mit Zeichnungen (Protagonist Peter möchte ein Puzzle lösen), um die untersuchten Kinder und Erwachsenen zum Hypothesenbilden und -prüfen aufzufordern.

(III) „Nachdenken über Experimentieren“:

*- Interviews:*

Carey, Evans, Honda, Jay und Unger (1989) setzten die Methode des Interviews ein, um Veränderungen durch einen dargebotenen Unterricht zu untersuchen. Sie befragten ihre Proband/innen zum Wesen/zur Gestalt von Wissenschaft, einer Hypothese und eines Experimentes sowie warum Wissenschaftler ihre Arbeit machen und was sie bei unerwarteten Ergebnissen tun.

Betrachtet man nun mit Blick auf die zu diagnostizierende Experimentierkompetenz, bei der das Planen, Durchführen und Auswerten eines Experimentes im Sinne eines Problemlöseprozesses zentral sind (Mayer, 2007), und mit Blick auf die bestehende Schulpraxis (Pädagogische Diagnostik) alle aufgeführten wissenschaftlichen Diagnoseverfahren, lassen sich sehr deutlich geeignete von weniger geeigneten Methoden abgrenzen.

Oft werden in allen Bereichen der Schulpraxis Paper-pencil-Testverfahren eingesetzt. Werden schriftliche Aufgaben (z. B. Multiple-Choice) oder das Arbeiten mit gedanklichen Experimenten zur Datenerhebung verwendet, wird die Sache (das Experimentieren) vereinfacht oder verfremdet, die erhobenen Daten können aus diesem Grund nicht ohne Weiteres mit realem Handeln beim Experimentieren gleichgesetzt werden. Schecker und Parchmann (2006) heben hervor, dass bei der Einschränkung auf kognitive Fertigkeiten und Fähigkeiten der Aspekt des praktisch, manuellen Ausführens vernachlässigt wird. Praktische und manuelle Fertigkeiten sind aber ein wichtiger Bestandteil des erkenntnisorientierten Experimentierens. Aus eben diesem Grund sind auch Verfahren aus der Realitätsdimension III „Nachdenken über Experimentieren“ nur bedingt geeignet, wenn man alle Aspekte (kognitive Fertigkeiten und Fähigkeiten sowie praktische und manuelle Fertigkeiten) einbeziehen möchte. Ein bestehender Unterschied zwischen den Verfahren wird auch durch Vergleichsuntersuchungen deutlich. Diese belegen, dass sich die Ausprägung der Schüler/innenleistung, wie sie durch einen Experimentiertest erhoben wird, von der durch schriftliche oder Computersimulationstests erhobenen Leistung unterscheidet (Hammann et al., 2007; Schreiber, 2012). Vergleichsuntersuchungen zwischen reinen Interviews und Realtests fehlen, trotzdem kann davon ausgegangen werden, dass Erhebungen mittels



Realtests die praktischen und manuellen Aspekte kongruenter abbilden. Nicht-Realtest-Formate eignen sich daher für die Schulpraxis nur zur Diagnose von kognitiven Fertigkeiten und Fähigkeiten sowie von Schüler/innenkonzepten über das Experimentieren als Methode. Vergleicht man nun alle angesprochenen Methoden der Datenerhebung bei Realtests (Realitätsdimension I “Reales Experimentieren”) unter dem Gesichtspunkt der Schulrealität mit Klassengrößen zwischen 25-30 Schüler/innen, wird schnell deutlich, dass nicht alle Methoden der Datenerhebung gleich gut geeignet sind. Die Datenerhebung mittels Videographie und anschließender Analyse wie auch die Erhebung und Auswertung nachgeschalteter Interviews sind für die alltägliche Praxis zu zeitintensiv und widersprechen der schulpraktischen Anforderung „Zeitökonomie“ (SA7). Zusätzlich zur Klassengröße muss auch noch einschränkend bedacht werden, dass jeder Lernende pro Schuljahr mehrmals diagnostiziert werden sollte, damit Leistungsänderungen früh und valide sichtbar werden. Will man genau und gültig Lernstandsveränderungen messen, dürfen die Messabstände nicht zu groß sein (Ingenkamp & Lissmann, 2008). Eine Verlaufsdiagnose in kleinen Abständen liefert Informationen, die gezielt für die Förderung und Rückmeldung eingesetzt werden können und dadurch die Lernmotivation erhöhen (ebd.). Unter dem Gesichtspunkt des zeitlichen Aufwands (SA7: „Zeitökonomie“) betrachtet, erscheinen während des Unterrichts durchführbare, kriteriengeleitete Beobachtungen von Schüler/innen beim Experimentieren möglich, wenn diese Art der Erhebung für wenige Schüler/innen z. B. zum Zwecke der Feindiagnostik eingesetzt wird. Hierfür werden aber gute, und vor allem für die Schulpraxis taugliche, Beobachtungsleitfäden benötigt, die objektives, reliables (SA1: „Erfüllung der Testgütekriterien“) und schnelles Beobachten (von unter Umständen nur Teilaspekten) des Experimentierens möglich machen. Im Bezug zur Erhebung einer Schüler/innenanzahl in der Größe einer Schulklasse muss darüber hinaus gewährleistet sein, dass ein Lehrer/eine Lehrerin ihrer Aufsichtspflicht gerecht werden kann. Diese Aufgabe erfordert die Aufmerksamkeit der Lehrkraft und limitiert sicherlich bei vielen Experimenten die Möglichkeit alle oder viele Schüler/innen einer Klasse kriteriengeleitet zu beobachten. Eine weitere Datenerhebungsmethode bei den Realtests ist das Auswerten eines Schüler/innen-Protokolls, welches während des Experimentierens vom Lernenden angefertigt wird und später von der Lehrkraft kriteriengeleitet ausgewertet werden kann. Diese Erhebungsmethode, die zwar ebenfalls die Realität nicht kongruent abbildet, erscheint in Kombination mit zielgerichteten, sehr kurzen Beobachtungen zum Umgang mit Laborgeräten im Hinblick auf kognitive Fertigkeiten und Fähigkeiten sowie praktische und manuelle Fertigkeiten als einzige für die Schulpraxis geeignet – alle schulpraktischen Anforderungen [SAs] würden nahezu erfüllt (SA1: „Erfüllung der Testgütekriterien“ mit Einschränkung wegen der Nutzung der Protokolle). Emden und Sumfleth (2012) finden zwischen Realtests und prozessorientierten Protokollen teils hohe Korrelationen, deren Zustandekommen sie auf einen Gewöhnungseffekt mit den Erhebungsmethoden zurückführen. Die Korrelationen fallen für Proband/innen an Gymnasien höher aus als an Gesamtschulen und sind zwischen einzelnen Aufgaben inkonsistent. Weitere Untersuchungen, um dies zu bestätigen, sind nötig. Zusammengefasst kann festgestellt werden, dass für die Schulpraxis (Pädagogische Diagnostik) in Hinblick auf kognitive Fertigkeiten und Fähigkeiten sowie praktische und manuelle Fertigkeiten eine Kombination aus Schüler/innen-Protokollen (angefertigt bei

Realexperimenten) mit sehr kurzen Beobachtungen zum Umgang mit Laborgeräten und eine Feindiagnose von Teilaspekten bei einzelnen Schüler/innen durch kriteriengeleitete ausgeprägte Beobachtung geeignet sind. Für die Diagnose von kognitiven und nicht praktischen, manuellen Aspekten sind Paper-pencil-Testverfahren geeignet (SA7: „Zeitökonomie“).

Die Aussagekraft eines formellen Diagnoseverfahrens ist aber nicht nur durch die Wahl der Erhebungsmethode, sondern ebenso durch die Wahl der diagnostischen Analyseeinheiten, die das zu messende Merkmal (in unserem Fall Experimentierkompetenz) abbilden, bestimmt. Die Auswahl diagnostischer Analyseeinheiten hat Einfluss auf die Aussagekraft und auf die Praxistauglichkeit des Diagnoseverfahrens. Beschränkte sich bisher der Blick auf SA1: „Erfüllung der Testgütekriterien“ und SA7: „Zeitökonomie“, treten nun die anderen SAs in den Blick.

#### **4 Welche Analyseeinheiten eignen sich für die Pädagogische Diagnostik?**

Wer den Lern- oder Leistungsstand eines Lernenden erhebt, sollte diesen mit geeigneten diagnostischen Analyseeinheiten, die unterschiedlichsten Anforderungen standhalten, erfassen. Wie zu Beginn bereits erwähnt, sind unter diagnostischen Analyseeinheiten Kategorien zu verstehen, auf deren Vorhanden- bzw. Nicht-Vorhandensein kriteriengeleitet geachtet wird (Beispiele, um Analyseeinheiten zu veranschaulichen: gemachte Fehler bei der Versuchsdurchführung, Kompetenzstufen bei der Hypothesenbildung etc.). Die Ausprägungen aller diagnostischen Analyseeinheiten ergeben das Diagnosebild eines Schülers/einer Schülerin. Auf dem Diagnosebild baut dann später die Beratung, der Förderplan oder die Beurteilung auf. Was sind aber die Anforderungen, die an diagnostische Analyseeinheiten in Bezug auf das erkenntnisorientierte Experimentieren gestellt werden müssen?

- (1) Das Vorhandensein der einzelnen diagnostischen Analyseeinheiten eines Analyseeinheitensystems muss deutlich und trennscharf erkennbar sein. Es darf nicht zu Unklarheiten, Verwechslungen oder Missverständnissen durch Überschneidungen kommen. Eine Analyseeinheit muss eindeutig einer Aussage (vorhanden, nicht-vorhanden) zugeordnet werden können (Teilaspekte von SA2: „Theoretische Präzisierung“; Hesse & Latzko, 2011).
- (2) Die diagnostischen Analyseeinheiten müssen eine geeignete Grundlage für Beratungsgespräche und Förderprogramme darstellen (vgl. hierzu auch: Hesse & Latzko, 2011; Ingenkamp & Lissmann, 2008).
  - Sie müssen Beschreibungen und Erklärungen des Leistungsstands möglich machen und sollten nicht nur Grundlage für subjektive Interpretationen bieten (SA6: „Vorhandene Beurteilungsmaßstäbe“, SA2: „Theoretische Präzisierung“). Beispielsweise ist die Aussage ‚Schülerin XY kann keine Hypothese aufstellen‘ ungeeignet, um individuelle Förderung abzuleiten. Die Aussage trifft auf mehrere potenzielle Lernschwierigkeiten der Schülerin zu: wenn sie keine Hypothese formulieren kann (sprachlich Schwierigkeiten hat), wenn sie nicht weiß, was eine



Hypothese ist (epistemologisches Verständnis) oder wenn sie keine Variablen selektieren kann. Möchte man später aus dem Diagnosebild einen Förderplan aufstellen, ist dies mit der sehr allgemeinen Feststellung unter Umständen nicht möglich, da keine exakt auf die Schwierigkeit angepassten Förderhinweise gegeben werden können.

- Zudem müssen die diagnostischen Analyseeinheiten auch für „Laien“ erkennbar und nachvollziehbar sein. Sie müssen nicht nur Lehrer/innen, sondern auch bei Schüler/innen und Eltern vorstellbare „Gebilde“ erzeugen, sodass individuelles, eigen- und fremdgeplantes Lernen möglich wird. Hierbei ist wichtig, dass die wissenschaftliche Beschreibung einer diagnostischen Analyseeinheit so konkret wie möglich ist (SA8: „Für Rückmeldung geeignet“).
- (3) Die diagnostischen Analyseeinheiten müssen das Erkennen auch geringer Lernstands- und Leistungsveränderungen ermöglichen, damit Lernerfolge schnell erkannt werden können und die Lernmotivation bei Schüler/innen angestoßen oder der Förderplan schnell angepasst werden kann. (SA3: „Lernstand sichtbar machen“, SA4: „Veränderungen sensibel erfassen“; Alderson, 2005; Walter, 2008).

Als Grundlage für diagnostische Analyseeinheiten bei der experimentellen Methode eignen sich Lernergebnisse (Feststellung von diesen) oder aber die Erfassung von Barrieren (Problempunkten), die Lernergebnisse verhindern. Problempunkte werden durch methodische Fehler, manuelle Fehler und Fehlvorstellungen angezeigt. Die Verwendung von Problempunkten erscheint für die Pädagogische Diagnostik gegenüber Leistungsfeststellungen weitaus geeigneter, da sie eine differenzierte Analyse von Hindernissen bei individuellen Lernprozessen ermöglichen. Mit einem Blick auf Lernergebnisse kann man im Diagnosebild hauptsächlich erkennen, welche Ausprägungen von Teilkompetenzen ein Schüler/eine Schülerin besitzt oder nicht besitzt (summatives Assessment). Für ein formatives Assessment müsste man nun im zweiten Schritt, um geeignete Beratung führen und Förderprogramme aufbauen zu können, untersuchen, welche genauen Ursachen (Problempunkte) beim Lernenden vorliegen, die das Defizit verursachen (Aspekt der Lernhindernisse bei SA3: „Lernstand sichtbar machen“). Verwendet man von Anfang an Problempunkte als diagnostische Analyseeinheiten, kann man direkt (ohne Umweg) pädagogisch wirksam werden, was eine große Zeitersparnis zur Folge hätte (SA7: „Zeitökonomie“). Aus den vorgefundenen und nicht vorgefundenen Problempunkten könnten sich zudem Kompetenzstufen ablesen lassen. Hierzu ist die Zuordnung auftretender Problempunkte zu den Kompetenzstufen eines bestehenden Kompetenzstufenmodells notwendig (z. B. dem Modell von Wellnitz et al., 2012). Mit der Verwendung von Problempunkten sind daher nicht nur formative, sondern bei Bedarf auch summative Assessments möglich.

## **5 Fazit**

Für die Pädagogische Diagnostik im Bereich Experimentieren stehen prinzipiell geeignete wissenschaftliche Erhebungsmethoden für kognitive Fertigkeiten und Fähigkeiten sowie

praktische und manuelle Fertigkeiten zur Verfügung (Schüler/innen-Protokolle und kriteriengeleitete Beobachtung). Da ein Diagnosebild über diagnostische Analyseeinheiten erzeugt wird, kann gute formelle Diagnose nur aufbauend auf aussagekräftigen diagnostischen Analyseeinheiten stattfinden. Ein notwendiger Theorierahmen, der eine Auflistung und Beschreibung von möglichen Lernschwierigkeiten sowie Vergleichswerten liefert, sollte aber noch weiterentwickelt werden. Es gibt bereits einige wissenschaftliche Arbeiten, die sich mit den Problembereichen und Problempunkten von Schüler/innen beim Experimentieren befassen (beispielsweise: Hammann, Phan, Ehmer & Bayrhuber, 2006; Kirchner & Priemer, 2007; Wahser & Sumfleth, 2008). Erfolgreiche Arbeiten berücksichtigen nur vereinzelt praktische, manuelle Fertigkeiten und Fähigkeiten. Langzeit- oder Querschnittsuntersuchungen wurden bisher nicht durchgeführt. Der Großteil der Erhebungen in der Sekundarstufe I fand in Gymnasien statt, die Schüler/innen der anderen Schularten und der Sekundarstufe II wurden selten untersucht. Weitere Arbeiten, die an den bisherigen ansetzen, diese fortführen und zusammenführen, wären wünschenswert. Für die Schulpraxis sinnvoll wäre hier auch die Einbindung von Fördermöglichkeiten in den Theorierahmen, was die Zusammen- und Weiterführung von bisherigen Interventionsstudien notwendig macht. Betrachtet man die Vermittlung von Experimentierkompetenz, wie sie in den nationalen Bildungsstandards verankert ist, unter dem Aspekt der individuellen Förderung, dann muss die noch nicht ausreichend geschlossene Lücke in der formellen Pädagogischen Diagnostik in diesem Bereich unbedingt geschlossen werden. Für Deutsch- oder Mathematiklehrer/innen ist es oft hilfreich und unabdingbar auf vorhandene formelle Diagnoseinstrumente zurückgreifen, um Defizite in der Lese- bzw. Rechtschreibkompetenz oder Defizite im mathematischen Bereich eines Schülers/einer Schülerin gesichert zu erkennen und geeignete Förderpläne hieraus abzuleiten. Sieht man Experimentierkompetenz als einen gleichermaßen wichtigen Bildungsbestandteil an, dann sollte solch eine Möglichkeit auch Naturwissenschaftslehrer / innen offen stehen.

## Dank

*Der Autor bedankt sich bei Juniorprof. Dr. Markus Emden, Benjamin Schmidt und Petra Georg für die Durchsicht des Manuskripts und für wertvolle Anregungen und Hinweise.*

## Literatur

- Abels, S. & Markic, S. (2013). Umgang mit Vielfalt - neue Perspektiven im Chemieunterricht. *Unterricht Chemie*, 24(135), 2-6.
- Alderson, J. C. (2005). *Diagnosing Foreign Language Proficiency: The Interface between Learning and Assessment*. New York: Continuum.
- Buholzer, A. & Kummer Wyss, A. (2010). *Alle gleich - alle unterschiedlich!. Zum Umgang mit Heterogenität in Schule und Unterricht*. Zug: Klett und Balmer.
- Carey, S., Evans, R., Honda, M., Jay, E. & Unger, C. (1989). 'An experiment is when you try it and see if it works'. A study of grade 7 students' understanding of the construction of scientific knowledge. *International Journal of Science Education*, 11(5), 514-529.

- Emden, M. & Sumfleth, E. (2012). Prozessorientierte Leistungsbewertung. *Der mathematische und naturwissenschaftliche Unterricht*, 65(2), 68-74.
- Hammann, M., Phan, T. T. H. & Bayrhuber, H. (2007). Experimentieren als Problemlösen: Lässt sich das SDDS-Modell nutzen, um unterschiedliche Dimensionen beim Experimentieren zu messen? *Zeitschrift für Erziehungswissenschaft*, 10(Sonderheft 8), 33-49.
- Hammann, M., Phan, T. T. H., Ehmer, M. & Bayrhuber, H. (2006). Fehlerfrei Experimentieren. *Der mathematische und naturwissenschaftliche Unterricht*, 59(5), 292-299.
- Hammann, M., Phan, T. T. H., Ehmer, M. & Grimm, T. (2008). Assessing pupils' skills in experimentation. *Journal of Biological Education*, 42(2), 66-72.
- Hesse, I. (2014). Pädagogisch-psychologische Diagnostik für Lehrkräfte - Herausforderungen, Aufgaben, Probleme. In A. Fischer, C. Höble, S. Jahnke-Klein, H. Kiper, M. Komorek, J. Michaelis et al. (Hrsg.), *Diagnostik für lernwirksamen Unterricht* (S. 15-39). Baltmannsweiler: Schneider.
- Hesse, I. & Latzko, B. (2011). *Diagnostik für Lehrkräfte* (2. Auflage). Regensburg: Barbara Budrich UTB.
- Huhta, A. (2008). Diagnostic and Formative Assessment. In B. Spolsky & F. M. Hult (Hrsg.), *The Handbook of Educational Linguistics* (S. 469-482). Malden (USA): Blackwell Publishing.
- Ingenkamp, K. & Lissmann, U. (2008). *Lehrbuch der Pädagogischen Diagnostik* (6. Auflage). Weinheim: Beltz.
- Kirchner, S. & Priemer, B. (2007). Probleme von Schülern mit offenen Experimentieraufgaben in Physik. In D. Höttecke (Hrsg.), *Naturwissenschaftlicher Unterricht im internationalen Vergleich* (S. 346-348). Münster: Lit Verlag.
- Klahr, D. (2000). *Exploring Science. The Cognition and Development of Discovery Processes*. Cambridge: MIT Press.
- Klauer, K. J. (1982). *Handbuch der Pädagogischen Diagnostik: Band 1*. Düsseldorf: Schwann.
- KMK (2009). *Empfehlungen der Kultusministerkonferenz zur Stärkung der mathematisch-naturwissenschaftlich-technischen Bildung (Beschluss der Kultusministerkonferenz vom 07.05.2009)*.
- KMK (2004). *Standards für die Lehrerbildung: Bildungswissenschaften (Beschluss der Kultusministerkonferenz vom 16.12.2004)*.
- Krüger, D. & Meyfarth, S. (2009). Binnen - kurzer Zeit - differenzieren! *Unterricht Biologie*, 347/348, 2-10.
- Langfeldt, H.-P. (2006). *Psychologie für die Schule*. Weinheim: Beltz.
- Langlet, J. (2009). *Individuelle Förderung im naturwissenschaftlichen Unterricht: Hilfe für die Umsetzung individuellen Kompetenzerwerbs* (MNU Themenreihe Bildungsstandards, 1. Auflage). Neuss: Klaus Seeberger.
- Mayer, J. (2007). Erkenntnisgewinnung als wissenschaftliches Problemlösen. In D. Krüger & H. Vogt (Hrsg.), *Theorien in der biologiedidaktischen Forschung: Ein Handbuch für Lehramtsstudenten und Doktoranden* (S.177-186). Berlin: Springer.
- Meier, M. & Mayer, J. (2012). Experimentierkompetenz praktisch erfassen. Entwicklung und Validierung eines anwendungsbezogenen Aufgabendesigns. In U. Harms & F. X. Bogner (Hrsg.), *Lehr- und Lernforschung in der Biologiedidaktik: Band 5* (S. 81-98). Innsbruck: Studien Verlag.
- Priemer, B. (2011). Was ist das Offene beim offenen Experimentieren? *Zeitschrift für Didaktik der Naturwissenschaften*, 17, 315-337.
- Pysik, A. & Bauer, C. (2013). Diagnostizieren und fördern - systematisch und effizient. Ein Differenzierungskonzept für den mathematisch-naturwissenschaftlichen Unterricht. *Praxis der Naturwissenschaften Physik in der Schule*, 62(6), 5-17.
- Schauble, L., Klopfer L. E. & Raghavan K. (1991). Students' Transition from an Engineering Model to a Science Model of Experimentation. *Journal of Research in Science Teaching*, 28(9), 859-882.
- Schecker, H. & Parchmann, I. (2006). Modellierung naturwissenschaftlicher Kompetenz. *Zeitschrift für Didaktik der Naturwissenschaften*, 12, 45-66.
- Schrader, F.-W. (2010). Diagnostische Kompetenz von Eltern und Lehrern. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (4., überarbeitete und erweiterte Auflage, S. 102-108). Weinheim: Beltz.
- Schreiber, N. (2012). *Diagnostik experimenteller Kompetenz*. Berlin: Logos.

- Schreiber, N., Theyßen, H. & Schecker, H. (2009). Experimentelle Kompetenz messen?! *Physik und Didaktik in Schule und Hochschule*, 8(3), 92-101.
- Schrempp, I. & Sodian, B. (1999): Wissenschaftliches Denken im Grundschulalter. Die Fähigkeit zur Hypothesenprüfung und Evidenzevaluation im Kontext der Attribution von Leistungsergebnissen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 31(2), 67-77.
- Schulz, A., Prinz, E. & Wirtz, M. (2012). Schüler planen Experimente und testen Hypothesen. Diagnose von Experimentierkompetenzen und mehrbenenanalytischer Klassenstufen- und Schulartvergleich. In W. Rieß, M. Wirtz, B. Barzel & A. Schulz (Hrsg.), *Experimentieren im mathematisch-naturwissenschaftlichen Unterricht: Schüler lernen wissenschaftlich denken und arbeiten* (S. 333-352). Münster: Waxmann.
- Tent, L. & Stelzl, I. (1993). *Pädagogisch-psychologische Diagnostik: Band 1 Theoretische und methodische Grundlagen*. Göttingen: Hogrefe.
- Wahser, I. & Sumfleth, E. (2008). Training experimenteller Arbeitsweisen zur Unterstützung kooperativer Kleingruppenarbeit im Fach Chemie. *Zeitschrift für Didaktik der Naturwissenschaften*, 14, 219-241.
- Walter, J. (2008). Curriculumbasiertes Messen (CBM) als lernprozessbegleitende Diagnostik: Erste deutschsprachige Ergebnisse zur Validität, Reliabilität und Veränderungssensibilität eines robusten Indikators zur Lernfortschrittsmessung beim Lesen. *Heilpädagogische Forschung*, 34(2), 62-79.
- Wellnitz, N., Fischer, H. E., Kauertz, A., Mayer, J., Neumann, I., Pant, H. A., Sumfleth, E. & Walpuski, M. (2012). Evaluation der Bildungsstandards - eine fächerübergreifende Testkonzeption für den Kompetenzbereich Erkenntnisgewinnung. *Zeitschrift für Didaktik der Naturwissenschaften*, 18, 261-291.

## **Kontakt**

Dr. Armin Baur

Pädagogische Hochschule Schwäbisch Gmünd, Institut für Naturwissenschaften, Abteilung Biologie, Oberbettringer Str. 200, 73525 Schwäbisch Gmünd, armin.baur@ph-gmuend.de